PROBABILIDADES Y ESTADÍSTICA (C)

Práctica 9

1. Una empresa vende dos variedades de soja. La variedad 1 tiene un rendimiento por ha. que puede considerarse una variable aleatoria con distribución $N(37, 25)$, y la variedad 2 tiene un rendimiento por ha. que puede considerarse $N(40, 25)$. Un cliente realizó una compra de semillas de la variedad 2 y antes de continuar comprando a esta empresa, quiere asegurarse de que las semillas que le enviaron realmente pertenecen a esa variedad y no a la variedad 1.

   Con ese fin, cultiva 10 parcelas de 1 ha. y obtiene los siguientes rendimientos:

   37 - 39.50 - 41.70 - 42 - 40 - 41.25 - 43 - 44.05 - 38 - 38.50

   El cliente quiere que la probabilidad de seguir comprando a esta empresa cuando las semillas no son de la variedad pedida sea 0.05.

   a) Explicar por qué las hipótesis para este problema son

   $$H_0 : \mu = 37 \qquad H_1 : \mu = 40.$$

   b) Encontrar un test para estas hipótesis y decir qué decisión se toma en base a la muestra dada. Calcular la probabilidad del error de tipo II.

   c) Encontrar el test para el caso en que se cultiven n parcelas (con igual nivel).

   d) Determinar el número n de parcelas a cultivar para que la probabilidad del error de tipo II sea menor o igual que 0.05.

   e) Explicar por qué el test planteado en c) sirve también para las hipótesis

   $$H_0 : \mu = 37 \qquad H_1 : \mu > 37.$$

   f) Calcular la función de potencia $\pi(\mu)$ del test planteado en c), verificar que es creciente para $\mu \in \mathcal{R}$ y deducir que este test sirve también para testear

   $$H_0 : \mu \leq 37 \qquad H_1 : \mu > 37.$$

2. En la construcción de un edificio debe usarse un concreto que tenga una tensión media mayor a 300 psi. Para verificar si el concreto preparado a partir del cemento "Loma Blanca" cumple con este requerimiento, se realizan 15 mediciones en forma independiente de la tensión de este concreto. Se observa una media muestral de 304 psi y un desvío estándar muestral de 10 psi. El constructor está dispuesto a correr un riesgo del 5% de comprar cemento "Loma Blanca" cuando éste produce un concreto que no cumple con las especificaciones. Suponiendo que los datos están normalmente distribuidos,

   a) Plantear un test de nivel 0.05 para testear las hipótesis

   $$H_0 : \mu = 300 \qquad H_1 : \mu > 300.$$

   ¿Qué decisión se toma?

b) Acotar el valor p.

3. Se diseñó un nuevo sistema de riego de manera tal que el desvío del tiempo de activación sea menor que 6 segundos. Se lo prueba 11 veces, obteniéndose los siguientes tiempos de activación:

   27 - 41 - 22 - 27 - 23 - 35 - 30 - 24 - 27 - 28 - 22

   Suponiendo que el tiempo de activación (en segundos) es una v.a. con distribución normal:

   (a) Testear a nivel 0.05 las siguientes hipótesis:

   $$H_0 : \sigma^2 \leq 36 \qquad H_1 : \sigma^2 > 36.$$

   ¿Usted decidiría que el sistema no cumple la especificación?

   (b) Acotar el valor p.

   (c) Encontrar la función de potencia $\pi(\sigma^2)$ y decir si es creciente o decreciente como función de σ^2.

4. Se supone que 1 de cada 10 fumadores prefiere la marca A. Después de una campaña publicitaria en cierta región de ventas, se entrevistó a 200 fumadores para determinar la efectividad de la campaña. El resultado de esta encuesta mostró que 26 personas preferían la marca A.

   a) Se desea deternminar si estos datos indican un aumento en la preferencia por la marca A. Para ello, plantear un test de nivel aproximado 0.05 las hipótesis:

   $$H_0 : p = 0.10 \qquad H_1 : p > 0.10.$$

   ¿Qué decisión se toma?

   b) Calcular el valor p.

   c) ¿Cuál es la probabilidad aproximada de decidir que la campaña publicitaria no fue efectiva, cuando en realidad la proporción de preferencia por la marca A después de la campaña es 0.15?

   d) ¿Qué tamaño de muestra debería tomarse para que la probabilidad de c) fuese a lo sumo 0.05?

5. Se supone que el tiempo de duración de cierto tipo de lamparitas tiene distribución exponencial de parámetro λ. Una fábrica garantiza que el tiempo medio de vida de las lamparitas que produce es mayor que 50 días, y la empresa vendedora quiere asegurarse que la producción satisface las especificaciones antes de sacarla a la venta. Para ello toma al azar una muestra de 15 lamparitas y observa el tiempo de duración de las mismas, obteniendo un promedio de 53 días.

   La empresa quiere tener un 95% de probabilidad de no vender si no se satisfacen los requerimientos, es decir, desea testear a nivel 0.05 las hipótesis:

   $$H_0 : \frac{1}{\lambda} = 50 \qquad H_1 : \frac{1}{\lambda} > 50.$$

a) Proponer un test de nivel exacto para este problema. ¿Qué decisión toma? (SUGERENCIA: Usar el item a) del Ejercicio 5 de la Práctica 8.)

b) Repetir a) usando un nivel aproximado.

c) Calcular la función de potencia aproximada para el test del ítem anterior. ¿Este test conserva el nivel aproximado 5% si ampliamos la hipótesis nula a $\frac{1}{\lambda} \leq 50$?

d) Utilizando el test aproximado, ¿qué probabilidad tiene la empresa de no sacar la producción a la venta, si el promedio de vida verdadero es 52 días?

e) ¿Cuál debería ser el tamaño de la muestra para que la probabilidad de c) fuese 0.1?

6. Se desea determinar si un dado de seis caras está cargado o no. Para ello se arroja el dado 1000 veces obteniéndose los resultados contenidos en el archivo **dado.txt**, que se encuentra en la página web de la materia.

   Estamos interesados en testear las hipótesis H_0 : El dado es equilibrado vs. H_1 : El dado está cargado. Para ello consideramos distintos tests:

   (a) Sea X = cantidad de veces que el resultado del dado es par. Plantear las hipótesis y determinar la región de rechazo para un test de nivel aproximado $\alpha = 0.05$ basado en X. Para el conjunto de datos obtenido, ¿cuál es el valor p aproximado?, ¿cuál es la decisión?.

   (b) Sea Y = cantidad de veces que el resultado del dado es menor o igual que 3. Plantear claramente las hipótesis y determinar la región de rechazo para un test de nivel aproximado $\alpha = 0.05$ basado en Y. Para el conjunto de datos obtenido, ¿cuál es el valor p aproximado?, ¿cuál es la decisión?.

   (c) Sea W = cantidad de veces que el resultado del dado es exactamente 3. Plantear claramente las hipótesis y determinar la región de rechazo para un test de nivel aproximado $\alpha = 0.05$ basado en W. Para el conjunto de datos obtenido, ¿cuál es el valor p aproximado?, ¿cuál es la decisión?.

   (d) Sea U = cantidad de veces que el resultado del dado es exactamente 2. Plantear claramente las hipótesis y determinar la región de rechazo para un test de nivel aproximado $\alpha = 0.05$ basado en U. Para el conjunto de datos obtenido, ¿cuál es el valor p aproximado?, ¿cuál es la decisión?.

   (e) En base a los resultados obtenidos. ¿Qué decisión tomaría? ¿Está cargado el dado?

7. En cada caso indique si la afirmación es verdadera o falsa y justifique:

   (a) El nivel de significación de un test es igual a la probabilidad de que la hipótesis nula sea cierta.

   (b) Un Error de tipo II es más grave que un Error de tipo I.

   (c) Si el p-valor es 0.3, el test correspondiente rechazará al nivel 0.01.

   (d) Si un test rechaza al nivel de significación 0.06, entonces el p-valor es menor o igual a 0.06.

(e) Si un intervalo de confianza de nivel 0.99 para la media μ de una distribución normal calculado a partir de una muestra da como resultado [-2.0, 3.0], entonces el test para las hipótesis $H_0 : \mu = -3$ vs $H_1 : \mu \neq -3$ basado en los mismos datos rechazará la hipótesis nula al nivel 0.01.